

<https://helda.helsinki.fi>

Investigating Proactive Search Support in Conversations

Andolina, Salvatore

ACM

2018-06-08

Andolina , S , Orso , V , Schneider , H , Klouche , K , Ruotsalo , T , Gamberini , L & Jacucci , G 2018 , Investigating Proactive Search Support in Conversations . in DIS '18 : Proceedings of the 2018 Designing Interactive Systems Conference . ACM , New York, NY, USA , pp. 1295-1307 , ACM conference on Designing Interactive Systems , Hong Kong , China , 09/06/2018 . <https://doi.org/10.1145/3196709.3196734>

<http://hdl.handle.net/10138/308884>

<https://doi.org/10.1145/3196709.3196734>

acceptedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Investigating Proactive Search Support in Conversations

Salvatore Andolina¹, Valeria Orso², Hendrik Schneider^{3,4}, Khalil Klouche³,
Tuukka Ruotsalo³, Luciano Gamberini², Giulio Jacucci^{1,3}

¹Helsinki Institute for Information Technology HIIT,

Department of Computer Science, Aalto University, Finland

²Human Inspired Technology Research Centre HIT, University of Padova, Italy

³Helsinki Institute for Information Technology HIIT,

Department of Computer Science, University of Helsinki, Finland

⁴University of Göttingen, Göttingen, Germany

ABSTRACT

Conversations among people involve solving disputes, building common ground, and reinforce mutual beliefs and assumptions. Conversations often require external information that can support these human activities. In this paper, we study how a spoken conversation can be supported by a proactive search agent that listens to the conversation, detects entities mentioned in the conversation, and proactively retrieves and presents information related to the conversation. A total of 24 participants (12 pairs) were involved in informal conversations, using either the proactive search agent or a control condition that did not support conversational analysis or proactive information retrieval. Data comprising transcripts, interaction logs, questionnaires, and interviews indicated that the proactive search agent effectively augmented the conversations, affected the conversations' topical structure, and reduced the need for explicit search activity. The findings also revealed key challenges in the design of proactive search systems that assist people in natural conversations.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords

Spoken conversation support; proactive search; voice interfaces; background speech.

INTRODUCTION

Casual collaborative tasks, such as travel planning or selecting a movie to watch create decision-making processes to solve disputes, build common ground, and reinforce mutual beliefs and assumptions based on people's preferences and prior knowledge. This process is conducted in a conversational

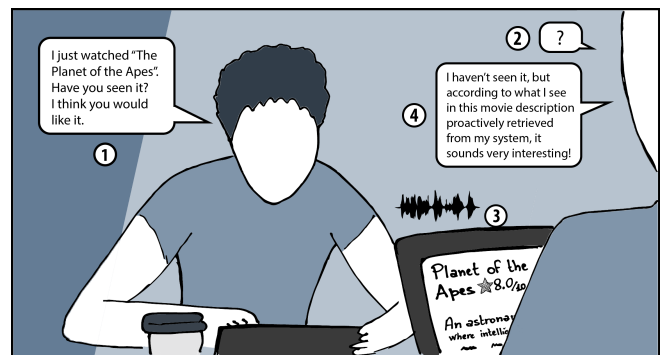


Figure 1. An illustration of the proactive search from spoken conversations. 1) The system listens to a natural spoken conversation between the participants; 2) information needs arise during the conversation; 3) the system is able to proactively retrieve useful information to support the conversation; 4) information needs are satisfied seamlessly and the conversation can continue smoothly.

exchange between two or more people, and a spoken form is often used to enable such an exchange. The key technologies assisting individuals in the decision-making processes are information retrieval and recommender systems, as well as question answering and summarization techniques [2, 34]. These systems can provide people with additional information to support their decisions and guide them to information that is important in the decision-making process.

Today, the availability of a wide array of personal devices has made search and recommendation possible in a variety of casual situations, including everyday conversations [9]. The ability to quickly conduct searches when co-located with others—a collaborative search—is believed to account for more than 60% of mobile searches [9, 45]. Often, however, search systems are not tailored to support the conversational process as such, but require explicit commands, preferences input, queries, and human attentional resources to guide the process [33]. As a consequence, the systems may disrupt social interactions rather than supporting them [1, 30, 32].

Proactive searches can leverage information from peoples' contexts to retrieve information in an easily accessible and non-intrusive manner [36]. Despite the current limitations

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

DIS '18, June 9–13, 2018, Hong Kong

© 2018 Copyright is held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 978-1-4503-5198-0/18/06...\$15.00

DOI: <https://doi.org/10.1145/3196709.3196734>

of automatic speech recognition, recent research on voice-based interaction [9, 27] has shown that relevant contextual information can be extracted even from a partial recognition.

These findings suggest that there are opportunities for scenarios in which a search is proactively performed in the background by using naturally occurring spoken conversational contexts between individuals while the individuals stay focused on their conversations and pay attention to their personal devices only when they need additional information. As opposed to previous approaches, which have mainly focused on supporting creativity [40], using voice or conversation recordings as research tools to uncover what users are doing with their mobile devices [27], or understanding information needs in conversations as parts of search tasks [41], we aim to go a step further by investigating how to proactively retrieve information from the Web to augment conversations (Figure 1).

In detail, *we aim to investigate how a proactive search agent can support natural spoken conversations between people by augmenting the conversations with additional information.* More specifically, we investigate the following research questions:

RQ1 Does a proactive search system with spoken input from a natural conversation influence the conversation?

RQ2 Does a proactive search system with spoken input from a natural conversation affect the consumption of Web resources during conversations?

RQ3 Does a proactive search system with spoken input from a natural conversation affect participants' subjective experiences?

To investigate our research questions, we designed SearchBot, an agent system that listens to a conversation, detects entities mentioned in the conversation, and proactively retrieves and presents information related to the conversation. We used SearchBot in a study with 24 participants (12 pairs) engaging in informal discussions on building travel or movie lists. Our findings suggest that SearchBot can effectively augment the conversations by enriching them with entities and documents shown on the screen and allowing people to consult the same number of Web resources as they would with traditional explicit searches. Subjective data from interviews and questionnaires suggest that proactive search support was generally found to be useful but also revealed key challenges for the design of proactive search systems assisting people in natural spoken conversation.

BACKGROUND

Information retrieval has traditionally relied on the query-response paradigm, with the user expressing information needs as explicit queries and the search engine responding with information items estimated to fulfill the user's information needs. Despite its practical success in Web search engines, the interactions between the user and the search engine may be laborious, as the broader search context, the user's exact search intent, and evolving information needs can be difficult to capture without explicit user involvement [48].

Proactive search refers to an information retrieval setting where the search system tries to automatically or proactively *anticipate* the user's upcoming queries and information needs [12, 16, 25]. An early attempt to build a proactive search system is the *Remembrance Agent* [35], which indexes a user's personal data such as emails and written notes. The system runs continuously in the background and displays a list of document summaries related to the current document being read or written. *Letizia* [24] is another early system that provides proactive recommendations during Web browsing using a set of heuristic rules. Another example of using search history is proposed in [42], where patterns repeated over time are extracted and used to proactively recommend resources to the user at specific times of the day. Recent work has also deployed computer vision to automatically detect broader user context and infer the user's work tasks for which supporting information is retrieved proactively [46, 47]. An interesting application for proactive search that comes close to our approach utilizes the subtitles of TV broadcasts being viewed by the user as context for predicting potential information needs [20].

Commercially deployed examples include *Google Now* and *Microsoft Cortana*, which run on users' smartphones and provide resources based on the current context. In particular, Google Now tries to model not only short-term search intents but long-term interests and habits based on several months of search log data [18].

A related research area is search personalization, where the search engine tries to discern individuals' unique search goals [43, 44]. Typical techniques include anticipating users' needs by taking into account their search histories [3], pre-search context information [23], social networks [11], and interaction behaviors [19, 37]. Researchers have also used previous search queries by the same user as context [10] and considered various kinds of relationships between the subsequent queries as features, such as reformulating the query or narrowing the search scope [49].

Supporting Conversations

Conversational systems have been studied from various angles, often referred to as situated interaction [8, 13]. Challenges of situated interactions include, for example, modeling initiatives in interaction, contextual interpretation, grounding, and turn-taking.

Related work has also investigated various kinds of visualizations to support conversations. *ConversationCluster* [7] uses visualizations to highlight salient moments in live conversation while archiving a meeting. Similarly, *Second Messenger* [14] uses a speech-recognition engine as an input method and shows filtered keywords from the group's conversation on an interactive display with the goal of increasing the visibility of diverse viewpoints. Other work [15] shows individual speaker-participation rates on a shared display to influence group behavior during a conversation.

Another main stream of research on conversation support systems has focused on creative design discussions. Schiavo et al. [39] introduced a system that monitors group members' non-verbal behaviors and promotes balanced participation,

giving the participants targeted directives through peripheral displays. *InspirationWall* leverages speech recognition and information exploration to augment a creative conversation with keywords that relate to the speech stream [4]. Similarly, *IdeaWall* [40] provides visual stimuli to the participants of a brainstorming session to facilitate the creative process. *Crowd-board* augments brainstorming conversations with real-time creative input from online crowds [6].

In this work, we aim to support a wider range of informal everyday conversations by augmenting them with information retrieved proactively by a search agent that listens to conversations.

Using Background Speech for Interaction

Speech-based interaction has been thoroughly studied in the literature. However, the interest in speech-based systems seems to have risen again in recent years, probably due to the recent advances in automatic speech recognition [31]. In particular, a large body of work focuses on a dialogic mode of interaction [28] where users communicate with the system using natural language. Commercially available examples include *Apple's Siri*, *Microsoft's Cortana*, and *Google Now*.

Less investigated is the use of background speech for interaction. One example is *Ambient Spotlight* [21], which uses speech recognition during meetings to search for desktop documents and puts them in a folder associated with the calendar entry related to that meeting. Other systems use background speech to retrieve words and other kinds of visual stimuli to support a creative conversation [4, 40]. As opposed to those systems, which are designed to support creative conversations where even misrecognitions and random results may lead to useful stimuli [22], we investigate how to support more generic conversations by proactively retrieving richer sources of information, such as documents, from the Web.

An important study related to our work is that of McGregor and Tang [26]. The aim of their study was to understand how well a speech-based agent could detect useful actions during formal meetings. Although the study used a simulated system to create a best-case scenario, results showed that extracted action items failed to fit with the meeting or gave an incorrect summary of what was being discussed or what the participants intended. A different approach was that of McMillan et al. [27]. Their study suggested that a continuous speech stream, rather than containing directly actionable items, can be used to identify users' next actions such as searches. This result inspired our research, as it means that regardless of the limitation of current automatic speech recognition technology, many useful words that would likely be used for a search could still be recognized. In this study, we aim to understand whether performing those searches proactively during conversations could effectively enrich those conversations.

THE PROACTIVE SEARCH SYSTEM

We designed the SearchBot system to monitor a conversation and provide continuous recommendations of related documents and entities in a non-intrusive way. Below, we describe

the system's main components: spoken conversation analysis, user interface design, and recommendation and retrieval methodology.

Spoken Conversation Analysis

SearchBot listens to conversations through a microphone. Speech recognition is performed by using Google's implementation of the HTML5 Web Speech API¹. The speech API takes an audio recording as an input and outputs a transcript in natural language. The speech recognizer is continuously listening to the conversation. The voice activity is automatically detected based on the audio input, and the system starts building a sentence from the input. After the activity stops, the system returns the recognized sentence. As soon as the system recognizes and returns the sentence transcript, it triggers the entity detection and recommendation component.

Entity Detection and Recommendation

Each transcription is processed by Google's Cloud Natural Language API², which is used to extract recognized entities from the transcripts. The API returns entities along with the information about their named entity types. For example, people, locations, and organizations are separately typed.

In order to recommend new entities based on the detected entities, we model them using a vector space model [38]. To recommend new entities, we train an entity embedding model using Word2Vec [29] on a complete English Wikipedia. The detected entities from the present transcript are each represented as a vector in the embedding space. The embedding model is used by first combining the vectors of the words in the recognized entities and then retrieving new entities by ranking other entities using their cosine similarity in the embedding space. Altogether, four highest-ranking entities are retrieved in response to each transcript (Figure 2c₂). For example, for the input "Bordeaux", "France", and "wines", the system computes a cosine distance for an input vector "Bordeaux" + "France" + "wines" and retrieves the entities "Bandol," "sauternes," "wines," and "Marseille," which have the smallest cosine distance to that vector.

Document Retrieval

Related documents are retrieved via Google Custom Search by combining entities recognized in the present transcript to a query. Entities of type "location" or "person" are prioritized to improve the relevance of the shown results. If an entity of such type is identified, a separate query is generated using that named entity, and the other entities are combined to that query. Altogether, four search results are retrieved in response to each transcript (Figure 2c₁).

More specifically, anytime the recognizer detects pauses, silence, or non-speech audio, a new sentence is returned. From the sentence, a set of entities is extracted, and a type is determined for each entity. If some of those entities are named entities, in our case of type "location" or "person," they are stored in a separate named entity query vector. All the entities are also used to form another general query vector. The final

¹<https://w3c.github.io/speech-api/speechapi.html>

²<https://cloud.google.com/natural-language/>

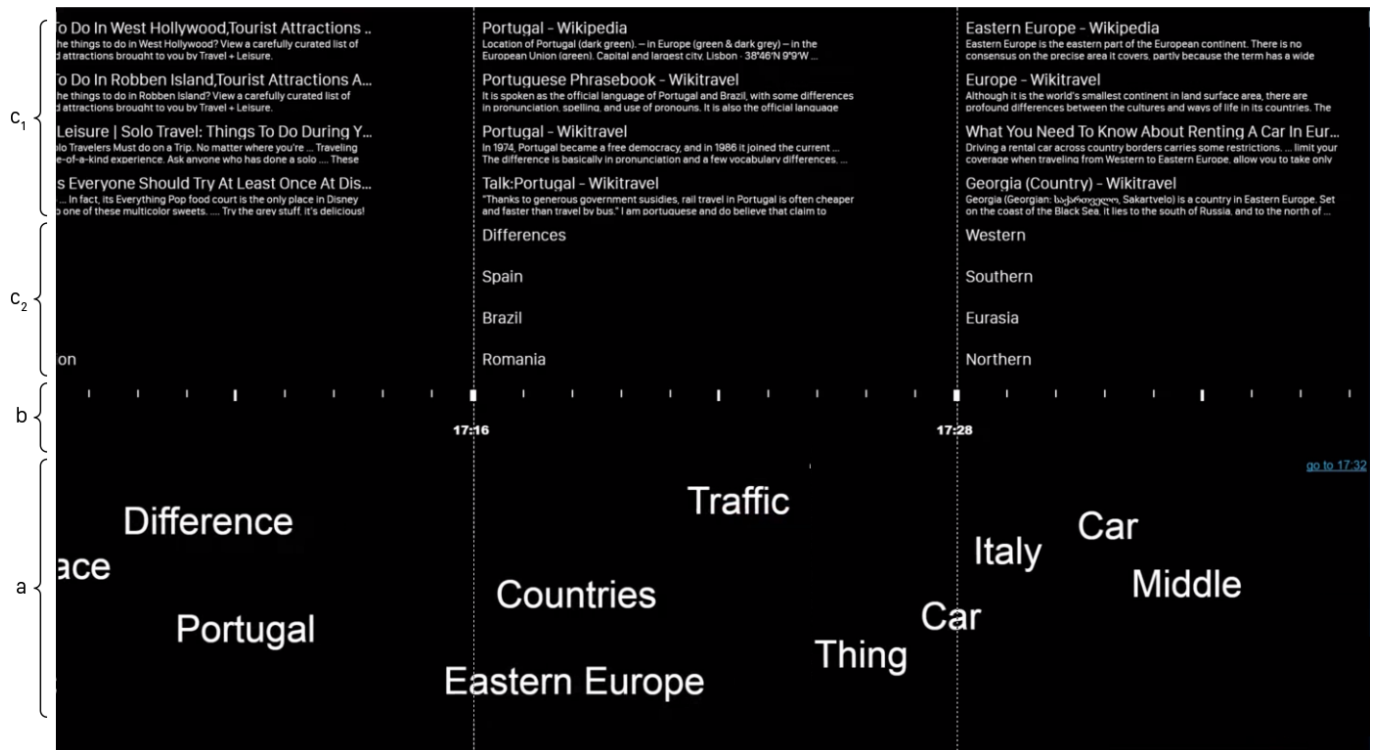


Figure 2. The user interface of the SearchBot system. The system monitors a conversation and provides continuous recommendations of related documents and entities in a non-intrusive way. a) Stream of recognized entities; b) timescale with timecodes; c₁) recommended documents; c₂) recommended entities.

set of search results shown to the user is then computed as the union of the highest ranked results in response to both query vectors. In case none of the entities are of type location or person, only the latter vector is used for retrieval.

An example of a sentence, extracted entities and their types, and the query vectors is given below:

```
SENTENCE: Bordeaux is famous for its wines.
ENTITIES: Bordeaux (type location), wines (type consumer good)
NAMED ENTITY QUERY VECTOR: Bordeaux
GENERAL QUERY VECTOR: Bordeaux + wines
```

User Interface Design

The user interface operated on a regular Web browser. It consists of a timeline that displays a stream of recognized entities in the lower part of the window (Figure 2a), a timescale with timecodes displayed in the center (Figure 2b), and successive sets of four retrieved documents (Figure 2c₁) and four recommended entities in the upper part of the window (Figure 2c₂). A new set extends the timeline every time a new transcription is available.

The user can interact with the system in multiple ways. Clicking on recognized or recommended entities triggers a search and opens the most relevant article in a new tab. Clicking on a document will open its content in a new tab. Users can

also move back and forth in the timeline by clicking on and dragging the central portion of the window.

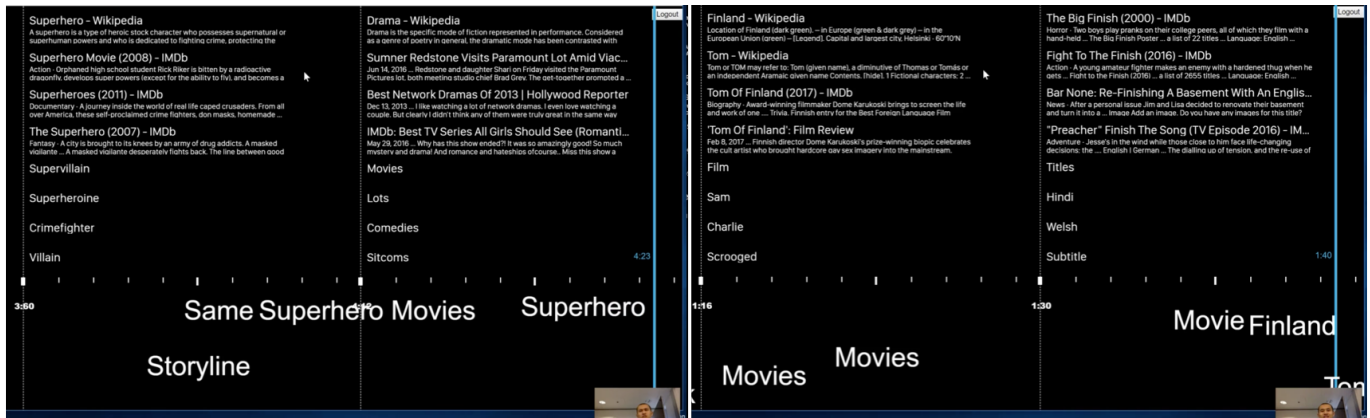
An example of system screen captures during a spoken conversation in which participants were having a natural conversation about movies is shown in Figure 3.

USER STUDY

A controlled laboratory experiment was designed to answer the research questions. SearchBot was originally conceived to support informal conversations, ideally occurring in any place. However, arranging the test in a natural environment (e.g., a cafeteria) would have subjected the system to a number of uncontrollable factors (e.g., ambient noise and incidental conversations) that could have influenced the system's performance and participants' experience. Since we did not know whether the items proposed by SearchBot could effectively feed a conversation, we chose to limit potential confounding factors by keeping the test in a controlled setting.

Experimental design

The experiment followed a within-subjects design with one independent variable being the system in use. Informal conversations on building travel or movie lists were supported either by SearchBot or by a traditional search engine used as the control. The order of presentation of systems and topics of discussion were counterbalanced across participants.



...
 Participant 1: Yeah. For superhero movies I was very into it because I just-- you know when I was little boy they were all, they were all around me...
 Participant 1: That was-- that was when I thought, "Wow, a movie can be made like this".
 Participant 1: Oh that's cool. The-- besides super hero movies, I'm also into dramas.
 ...

...
 Participant 2: What's the last movie you saw in a theater?
 Participant 1: [sighs] The last movie? Can-- Oh yeah, the last time I went to theater was for the movie Tom of Finland actually.
 Participant 2: Oh.
 I gotta see that. Probably should watch it. Was it good?
 ...

Figure 3. Two example screen captures of the system in a session where the participants were having a natural conversation about movies. The corresponding transcripts of the spoken conversations are shown below the screen captures. The system is recognizing and recommending entities and matching documents based on the conversational input.

Materials

We asked participants to complete three questionnaires. The first questionnaire was meant to collect background information (first and last name, age, provenance, education) and expertise in the field of HCI and previous experience with conversational agents. We also asked participants to assess the quality of the entities displayed by the system. More specifically, after they had used the experimental system, they were shown a list of the 100 suggested entities most frequently displayed during the conversation. For each word, we asked them to indicate if the word was pertinent to the conversation (namely, relevant) and if it was effectively mentioned in the conversation (namely, mentioned). A third questionnaire aimed at investigating their experience with the system in use. More specifically, we devised the questionnaire ad hoc, and it consisted of 12 items, exploring the impression that the system affected the conversation (items 1-4), the quality of the experience with the system (items 5-8), and the perceived quality of the entities shown (items 9-12). Participants had to indicate their level of agreement with each item on a 5-point Likert scale. For items 9-12, the option "not applicable" was also available. We asked participants in the experimental condition to complete all three questionnaires, and we asked participants in the control condition to complete only the third one. Finally, we devised a semi-structured interview to capture participants' direct comments and impressions of the system. More specifically, during the interview, we asked them to report their overall impressions of the system and the entities displayed. Furthermore, they had to comment on whether they had the impression that the system affected the conversation and whether they got interesting information. After the second session, we asked them to compare their experiences with both systems.

Equipment

For the present experiment, each participant used a MacBook Pro 15" laptop connected to a Samson Meteor microphone. The experimental session was video-recorded using a Panasonic camcorder. Additionally, the screen recording was taken using Screencast-O-Matic software, which also recorded participants' faces with the webcam embedded in the laptop.

Setting

The experiment took place in the laboratory. We set up the room to resemble a comfortable and informal environment, where participants could feel at ease. They sat at a desk in front of each other. Each participant had a laptop in front



Figure 4. Experimental setup. Participants were sitting around a table, and a laptop was placed in front of each participant. The laptops were displaying the SearchBot interface. Microphones were placed on the table to record the conversation.

of him/her. The laptop allowed them to easily maintain eye contact with their conversation partners and quickly glance at the screen (Figure 4).

Procedure

The experimental procedure consisted of two main phases, each corresponding to the system in use. On the day of the test, participants were first welcomed by the experimenter, who introduced them to the experiment’s main goals and overall procedure. After that, participants signed the informed consent. Phase 1 started with a training session on how to use the system; when everything was clear, the experimental session began. During the experiment, the experimenter simply asked participants to talk with their partners for 20 minutes. The experimenter then left the room to allow participants to talk freely. He followed the experimental session through a video connection and was reachable in case participants needed assistance. The task assigned was not meant to generate a specific outcome; rather, it was intended to provide only a general shape to the conversation. More specifically, we asked participants to share their experiences regarding the movies or travels (depending on the experimental condition) that had impressed them and to get inspirations from their partners’ words. We did not force the participants to use the system, but we allowed them to freely utilize or ignore recommendations according to their needs. After 20 minutes, the experimenter returned to the laboratory and asked participants to complete the online questionnaire about the quality of the entities shown. Next, the experimenter accompanied one participant out of the room to complete the post-experience questionnaire while the other participant remained in the laboratory and completed the semi-structured interview. After they both finished, they swapped places. Phase 2 unfolded exactly as Phase 1 did, with the only exception being the system in use. We used the same instructions in both experimental conditions. In both cases, we left participants free to use or ignore the system according to their needs.

Control system

We used Google Custom Search to create a custom search engine that would mimic the behavior of the APIs used in the experimental condition, while maintaining interaction and the look and feel typical of traditional search engines (Figure 5).

Datasets

To provide more relevant results for our tasks, we set Google Custom Search so that it would emphasize selected websites regarding movies or travels, as well as Wikipedia. We applied the same setting to both experimental conditions, and we restricted the search engine to specified domains³.

³We used the following domains in the movie task: www.hollywoodreporter.com, www.imdb.com, www.themoviedb.org, and www.rottentomatoes.com. We used the following domains in the travel task: www.wikitravel.org, www.travelandleisure.com, www.worldtravelguide.net, and www.tripexpert.com. In addition, we used the English Wikipedia for both tasks.

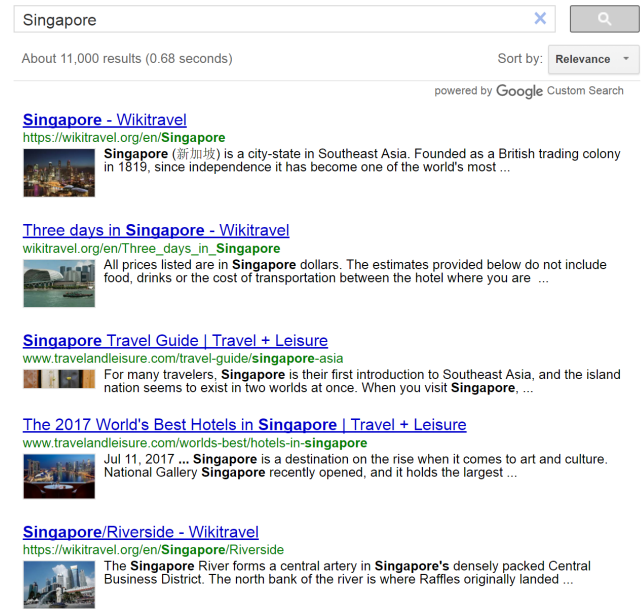


Figure 5. The instance of Google Custom Search used as a control.

Participants

A total of 24 participants (12 female) took part in the present study. The participants’ mean age was 27 years ($SD = 3.87$). Of the participants, 12 were undergraduate students, five were doctoral students, three were research assistants, three were post-doc researchers, and one was a nurse. Overall, 11 participants reported having previous experience with conversational agents, and all of them reported rare usage of them. They received two movie tickets as compensation for participating in the experiment.

Measures

In order to assess the the proactive search agent’ effect on the conversation, we used objective and subjective measures.

Influence of information shown on the conversation. To understand whether the information the proactive search agent presented on the screen influenced the conversation, we counted the entities extracted from the items shown on the screen that were mentioned in the 60 seconds following their first appearance on the screen. To control for possible cases in which displayed entities were mentioned by chance, we performed the same calculation in the control condition. In this case, the proactive search interface was running in the background, and results were not shown to the participants. To perform this analysis, we used a script on system logs and transcripts of the conversations obtained through a professional service.

Consumption of web resources. The number of pages opened by participants during the conversation served as a proxy for the consumption of Web resources. More specifically, the research prototype logged the pages accessed in the experimental condition, and for the control condition, the pages were traced through the navigation history of the Web browser used.

Perceived quality of the recommended items. We showed participants the list of the 100 recommended entities and the list of the 100 Web documents that the system displayed most frequently, and we asked them to mark the items that they considered pertinent and relevant to the conversation. We considered this measure a proxy for the perceived quality of the items the system suggested.

Preferred items with the proactive search agent. In the experimental condition, we logged the item types (i.e., Web documents, recommended entities, and recognized entities) that the system displayed and that the user clicked on to seek more information.

Subjective experience. We investigated participants' subjective experience with the system using a questionnaire and a semi-structured interview. We devised the questionnaire *ad hoc*, and it explored aspects related to the participants' impressions that the system had affected the conversation, the quality of the experience of use, and the overall relevance of the items displayed. The average score for each dimension was computed and then compared between the two groups. Concerning interviews, we transcribed participants' answers and ran a thematic analysis of those answers. We reviewed transcripts, identified recurring themes, and organized them into a codebook. We then applied the codes to the corpus of data [17].

FINDINGS

Influence of the information shown on the conversation

A Wilcoxon Signed-Ranks Test indicated that the entities extracted from the items shown on the screen were effectively mentioned in the conversation. More specifically, the number of extracted entities that participants mentioned in their conversations was significantly higher in the experimental condition (in which the items were actually shown) than in the control (in which the system was running in the background): $z = 2.33$, $p = .02$ (exp. $M = 7.46$, $SD = 4.90$, $Mdn = 6.00$; con. $M = 4.37$, $SD = 2.20$, $Mdn = 4.50$). This finding indicates that the references to the entities shown on the screen in the experimental condition were not due to chance (Figure 6).

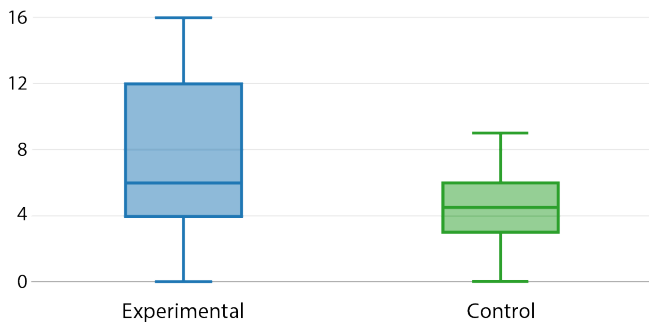


Figure 6. Number of entities that were mentioned within 60 seconds of when they were shown on the screen.

Consumption of Web resources

The consumption of Web resources, measured as the average number of Webpages opened, was also compared between the two conditions using a Wilcoxon test. The analysis did not highlight a statistical significant difference: $z = .4$, $p = .68$ (exp. $M = 11.75$, $SD = 10.01$, $Mdn = 11$; con. $M = 12.58$, $SD = 9.81$, $Mdn = 9.5$). This finding suggests that the system in use did not alter participants' search behavior.

Perceived quality of the recommended items

31.58% ($M = 31.58$, $SD = 14.70$) of the selected recommended entities were rated as relevant in the experimental condition, and the portion of the selected recommended documents rated as relevant was 17.38% ($M = 17.38$, $SD = 10.95$).

Preferred items with the proactive search agent

Figure 7 shows how participants clicked on the various types of items in the experimental condition. A non-parametric Friedman test among the various types of items clicked rendered a Chi-square value of 18.1, which was statistically significant ($p < .001$). Wilcoxon Signed-Ranks Tests with Bonferroni correction indicated that participants clicked on items representing Web documents significantly more than either suggested entities ($p < .01$) or recognized entities ($p < .01$), but there was no difference between the number of clicks on suggested entities and recognized entities ($p = .17$).

Questionnaires

We computed the average score for each dimension that the questionnaire assessed and compared the experimental and control conditions using a Wilcoxon test (Table 1). The analysis showed that participants had the impression that their conversations were affected to a greater extent by the system in the experimental condition than in the control condition: $z = 2.32$, $p = .02$ (exp. $M = 3.62$, $SD = .64$, $Mdn = 3.75$; con. $M = 3.15$, $SD = .86$, $Mdn = 3.00$). Similarly, the reported quality of the experience of use was more positive for the experimental condition ($M = 3.55$, $SD = .75$, $Mdn = 3.50$) than for the control condition ($M = 3.14$, $SD = .76$, $Mdn = 3.37$), $z = 2.01$, $p = .03$. No difference emerged between the conditions regarding the perceived relevance of the entities shown $z = .59$, $p = .55$ (exp. $M = 3.5$, $SD = .57$, $Mdn = 3.62$; con. $M = 3.4$, $SD = .51$, $Mdn = 3.5$).

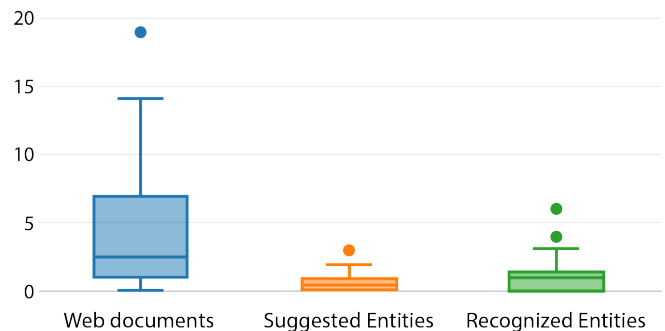


Figure 7. The types of items participants clicked on in the experimental condition. The y-axis is the number of clicks per session.

Table 1. Questionnaires

	SearchBot			Control			Comparison
	M	SD	Mdn	M	SD	Mdn	Wilc. Test
Influence on the conversation							
1. The system can give the conversation new directions	3.88	0.74	4.00	3.00	1.06	3.00	$z = 2.32$ $p = .02$
2. The system provides little support for enriching the conversation	2.88	1.19	3.00	3.08	1.21	3.00	
3. The system has the potential to influence what people are about to say	3.88	0.80	4.00	3.13	1.08	4.00	
4. The system had almost no effect on the conversation	2.38	0.88	2.00	2.46	1.14	2.00	
Quality of the experience							
5. The system was frustrating	2.25	0.90	2.00	2.46	1.10	2.00	$z = 2.01$ $p = .03$
6. Using the system was fun	3.67	0.87	4.00	2.88	0.85	3.00	
7. Using the system was effortful	2.79	1.10	3.00	2.79	1.14	3.00	
8. Using the system was pleasant	3.58	1.02	4.00	2.96	0.81	3.00	
Relevance of the items shown							
9. The items shown were overall interesting in the conversation	3.67	1.01	4.00	3.38	1.01	4.00	$z = .59$ $p = .55$
10. The items shown were not relevant to the conversation	2.33	0.70	2.00	2.25	0.68	2.00	
11. The items shown were overall pertinent to the conversation	3.61	0.72	4.00	3.48	0.73	4.00	
12. The items shown were redundant	2.92	0.78	3.00	2.96	0.95	3.00	

Interviews

Overall, SearchBot was well-received by participants, with the majority of users reporting a positive experience ($N = 16$): “It was fun. . . . It’s different from everything I’ve tried before” (P7A). As interviews indicated, participants preferred SearchBot over the control condition ($N = 14$) because it allowed participants to maintain eye contact during conversation and provided users with information effortlessly:

I think the [experimental] system is much more useful and easier because with the [control] system, you have to make a decision—“Ok, I need to google something”—and here [experimental system], it’s just a big flow, and you need to watch if something comes up. If nothing comes up, you just ignore it (P4B).

The first one [experimental] helped you to keep the eye contact. . . . You don’t need to concentrate. . . . it gives you ideas sometimes before they come to your mind (P5A).

I think this one [experimental] was better. . . . It was more useful. It didn’t get me frustrated with wrong information, and even if it displayed the kind of stuff that were not relevant, it didn’t bother my eyes. [When something relevant was shown, I thought] this is relevant to our discussion; let’s click it. So it was better than the other one (P8B).

However, some users ($N = 7$) did not convey a clear preference between the two systems although they remarked that they were different. They highlighted that the control system was more convenient to use when they explicitly wanted to look up some specific pieces of information but that using SearchBot was easier because it required no input from the participants:

The [experimental] one is easier because you don’t have to do anything; you just say the names, and you got the links. But again, the [control] one was nice because you could use it to Google (P10B).

A smaller proportion of participants ($N = 3$) expressed a preference for the control system because it made them feel in control:

[With] the [control] system, [it] was a better experience maybe because I’m used to using a search engine in my daily life all the time during a conversation (P12A).

Notably, the majority of the participants ($N = 15$) did have the feeling that SearchBot affected the conversation either by offering the chance to deepen the current topic ($N = 8$) or by inspiring new points to discuss ($N = 7$).

It [the experimental system] gave us new information. . . . Like, we were unsure where the Red Square was, if it was Moscow or St. Petersburg, and we found that out (P7B).

It [the experimental system] supported the conversation as it went. For example, we were talking about Mongolia, and I think it suggested Genghis Khan, and of course Genghis Khan is part of Mongolia (P10A).

We didn’t know what to talk about next, and I looked at the [experimental] system, and it said, “Berlin,” so I was interested in that place, so I asked her if she had been there before, so it changed the direction (P3B).

However, seven users did not have the impression that the experimental system was a support for the conversation, either because they felt proficient with the topic (“I don’t have that much the impression of that [. . .] it was a topic in which I didn’t really need help [. . .] the topic was quite familiar, especially since we talked about places I have been” (P6A)) or because they could not find the entities suggested by the system in a timely manner (“Not much, to be honest . . . but I was able to see the potential of it . . . because it didn’t catch up with what we were talking about” (P12A))

The majority of the participants ($N = 14$) did not find the information provided unique because they said they could have found the same data using the search engines they usually employ: “I don’t think so because I could have Googled everything it was showing” (P1B). Nevertheless, nine participants thought that the entities the system displayed were not obvious:

I think the [experimental] system gave me a lot of unexpected results. So, without the system, I would not search for those keywords (P3B).

SearchBot supported a better understanding of the conversations' topics, according to 16 users: *"'Cause I can quickly check something. When we were not sure what the real answer was, we could just click and check" (P3A).* However, six participants did not have the same impression because of their prior knowledge of the conversation's topic: *"Maybe not because the topic was so easy, you don't need help with this topic because you know it" (P5B).*

The entities displayed by the system were generally considered either relevant or useful ($N = 14$):

We talked about the recent Star Wars movie, and the system displayed links to the relevant pages. We were discussing some details, and we were able to check them from the pages that were proposed (P8A).

They [the recommendations] were pretty good ... When you say something, you have four to five different choices, and I think they were pretty accurate. Sometimes even you say something not so important, you would find something important. For example, we talked about cheap, how cheap is gonna be, and instantly the system gave us, like, cheap flights (P7A).

When it was relevant, it was really useful because you just click on it and then find more information, for example the year [of the movie] or something like that (P9A).

However, three users found the entities not relevant for the conversation and thus distracting: *"Something totally not relevant appeared, which made me less concentrated on the conversation" (P9B).* Two participants commented that the search based on the combination of two entities didn't work for them: *"I was talking about how to go to Bordeaux by train, so it was [the system suggested] train and Bordeaux separately, but it would be much more useful if it was like 'train to Bordeaux'" (P6A).* Three participants complained that the entities were displayed too late with respect to the stream of conversation: *"Maybe it showed it [the entities] a little bit afterward ... but it was interesting information" (P7B).* Finally, only one participant reported that the entities were useless to the conversation.

DISCUSSION

Information spaces grow in size and richness, and users increasingly prefer information to be delivered to them proactively as a part of secondary tasks supporting their primary tasks. Consequently, conventional search interfaces fall short in allowing users to concentrate on their primary tasks, and supporting information access by anticipating users' needs has become a major bottleneck in many complex tasks.

We studied how proactive searches conducted by using input directly from natural conversations between individuals can support the conversations. We designed the SearchBot system and used it in an experiment to study the influence, the number of consumed resources, and the effect on the user experience of a proactive search interface in supporting conversations.

Answers to the Research Questions

Here we reflect on the research questions that we defined earlier.

RQ1: Does proactive search with spoken input from natural conversation influence the conversation? Yes. Figure 6 shows how participants in the experimental condition frequently referred to the entities and documents shown on the screen during their conversations. The comparison with the control condition, in particular, demonstrated that these references were not due to chance. This result indicates that not only did the proactive search system retrieve useful information, but the displayed information influenced the conversation, as questionnaires and interviews further confirmed.

RQ2: Does proactive search with spoken input from natural conversation affect the consumption of Web resources during conversations? No. There was no significant difference in the number of Web resources consulted between the experimental and control systems. This result suggests that participants retrieved the same number of useful resources supporting the conversation in both experimental conditions. However, while in one case the resources were automatically retrieved by the proactive search agent, in the other case, explicit query formulation and refinement was needed.

RQ3: Does proactive search with spoken input from natural conversation affect participants' subjective experience? Yes. In general, the reported quality of the experience of using the system was more positive for the experimental condition, as it allowed participants to keep eye contact with each other, enabling more fluent conversation. Participants reported that SearchBot allowed them to check facts and build common ground without needing to exert much mental effort. Furthermore, the system was able to expand the conversation in new directions. However, the added value of the proactive search experience seemed to come with the cost of feeling less in control of the search process. All in all, the participants were more satisfied with the SearchBot system.

Design Implications

In this section, we start from the lessons we learned in our study and discuss design implications to help set the stage for future developments of proactive search interfaces for conversation support.

Relevance of recommendations in conversation support.

Our study proved that errors in recognition and the consequent display of non-relevant results do not necessarily prevent a proactive search system from effectively supporting a conversation. Most participants were able to easily ignore non-relevant recommendations while benefiting from the relevant ones. However, while users can easily skim through a screen full of non-relevant results to identify a single piece of relevant information, it is important that at least that single piece of relevant information is there when needed. When this didn't happen, participants experienced distraction, frustration, and loss of trust in the system. Some heuristics may be required to improve the relevance of displayed results. In our case, we mostly used queries involving all the entities contained in a recognized sentence. However, this did not always produce the

desired results. Therefore, we used knowledge gained from pilot experiments to prioritize locations and people as the entities particularly relevant to our tasks. When such entities were detected, we used them in single-entity queries and showed the results on top of the list. This generally improved the system's capability to support the conversations, but it came at the cost of showing fewer recommendations deriving from combined entities. Deciding how to combine search terms and if and how to prioritize special terms is a key aspect to be considered when designing systems for proactive search support in conversations.

Types and number of recommended items. Our findings suggest that with SearchBot, around one third of recommended entities and around 17% of recommended documents were relevant to the conversation. Nevertheless, our findings also show that the most used items were Web documents. While in this initial investigation we chose to display the same number of recommended entities and documents, our findings suggest that future implementations should carefully consider how to allocate the screen's real estate for various kinds of items according to the type of discussion to be supported. In our study, participants did not use recommended entities much. They used them mostly to expand the conversation with new ideas. While this result confirms past research on creative conversations [4], it also highlights the fact that various item types are needed in conversations that are not merely creative, such as those explored in our study. Our findings suggest that richer sources of information, such as Web documents, should be given more importance in those cases, as they allow participants to check facts and build a common understanding on the topic of discussion.

Combining proactive and explicit searches. Our study suggests that proactive searches performed automatically by the system by using content extracted from spoken conversation allowed users to more easily maintain eye contact and stay focused on the conversation, as the interviews indicated. Most people preferred the proactive search approach to the explicit search one, but this preference came at the cost of losing control over the system. As the two approaches showed complementary strengths, next developments should consider integrating both modes of operations in the same interface.

Limitations and Future Work

While our work shows that proactive search support in conversation is already possible and provides several advantages over relying on traditional explicit search, it also has some limitations. Even if the system was devised to support informal and casual conversation, we chose to test it in a controlled setting. While this arrangement allowed us to control for confounding factors, it also limited our findings' ecological validity. Further research is needed to understand how proactive search support can affect conversation in more natural settings. Also, in this work, we designed a prototype with a limited set of features. The prototype served as a research tool to study the potential of proactive searches from natural spoken input to support conversations. This strategy allowed us to better understand the different features' roles. However, to understand the real potential of proactive search support in conversations, future

implementations should consider integrating more sophisticated features, such as interactions with the system to build and maintain intent models [5, 37] and using topic modeling to extract relevant context and improve the relevance of retrieved results.

CONCLUSION

The approach to proactive search used in this work utilizes the subtle human feedback signals observed directly from natural conversations, as opposed to previous work, which has mainly relied on conventional user input, such as issued queries or visited documents. We investigated how a proactive search agent that uses vocal conversational input could support informal conversations on travel or movie lists. We designed SearchBot, a proactive search agent that listens to conversations, detects entities mentioned in the conversations, and proactively retrieves and presents information related to the conversations. We used SearchBot in a comparative study with 12 pairs of participants. Our findings showed that information retrieved proactively by an agent listening to the conversation had the potential to effectively support the conversation with facts and ideas without causing much interruption to the conversation's flow but at the cost of participants feeling less in control of the search process. Findings also show that the proactive search approach retrieved the same number of useful resources supporting the conversation but without the participants needing to formulate explicit queries. Notably, this study allowed us to explore the design space of proactive search support in conversations, providing key design implications for the paradigm's future developments.

ACKNOWLEDGMENTS

This research was partially funded by the Academy of Finland (286440, 312274, 313610).

REFERENCES

1. Jesper Aagaard. 2016. Mobile devices, interaction, and distraction: a qualitative exploration of absent presence. *AI & SOCIETY* 31, 2 (01 May 2016), 223–231. DOI : <http://dx.doi.org/10.1007/s00146-015-0638-z>
2. Gediminas Adomavicius and Alexander Tuzhilin. 2005. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Trans. on Knowl. and Data Eng.* 17, 6 (June 2005), 734–749. DOI : <http://dx.doi.org/10.1109/TKDE.2005.99>
3. Eugene Agichtein, Eric Brill, and Susan Dumais. 2006. Improving Web Search Ranking by Incorporating User Behavior Information. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '06)*. ACM, New York, NY, USA, 19–26. DOI : <http://dx.doi.org/10.1145/1148170.1148177>
4. Salvatore Andolina, Khalil Klouche, Diogo Cabral, Tuukka Ruotsalo, and Giulio Jacucci. 2015a. InspirationWall: Supporting Idea Generation Through Automatic Information Exploration. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and*

- Cognition (C&C '15)*. ACM, New York, NY, USA, 103–106. DOI :
<http://dx.doi.org/10.1145/2757226.2757252>
5. Salvatore Andolina, Khalil Klouche, Jaakko Peltonen, Mohammad Hoque, Tuukka Ruotsalo, Diogo Cabral, Arto Klami, Dorota Glowacka, Patrik Florén, and Giulio Jacucci. 2015b. IntentStreams: Smart Parallel Search Streams for Branching Exploratory Search. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*. ACM, 300–305.
<http://doi.acm.org/10.1145/2678025.2701401>
 6. Salvatore Andolina, Hendrik Schneider, Joel Chan, Khalil Klouche, Giulio Jacucci, and Steven Dow. 2017. Crowdboard: Augmenting In-Person Idea Generation with Real-Time Crowds. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition (C&C '17)*. ACM, New York, NY, USA, 106–118. DOI :
<http://dx.doi.org/10.1145/3059454.3059477>
 7. Tony Bergstrom and Karrie Karahalios. 2009. Conversation Clusters: Grouping Conversation Topics Through Human-computer Dialog. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 2349–2352. DOI :
<http://dx.doi.org/10.1145/1518701.1519060>
 8. Dan Bohus and Eric Horvitz. 2009. Dialog in the Open World: Platform and Applications. In *Proceedings of the 2009 International Conference on Multimodal Interfaces (ICMI-MLMI '09)*. ACM, New York, NY, USA, 31–38. DOI :
<http://dx.doi.org/10.1145/1647314.1647323>
 9. Barry Brown, Moira McGregor, and Donald McMillan. 2015. Searchable Objects: Search in Everyday Conversation. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM, New York, NY, USA, 508–517. DOI :
<http://dx.doi.org/10.1145/2675133.2675206>
 10. Huanhuan Cao, Derek Hao Hu, Dou Shen, Daxin Jiang, Jian-Tao Sun, Enhong Chen, and Qiang Yang. 2009. Context-aware Query Classification. In *Proceedings of the 32Nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '09)*. ACM, New York, NY, USA, 3–10. DOI :
<http://dx.doi.org/10.1145/1571941.1571945>
 11. David Carmel, Naama Zwerdling, Ido Guy, Shila Ofek-Koifman, Nadav Har'el, Inbal Ronen, Erel Uziel, Sivan Yogeve, and Sergey Chernov. 2009. Personalized Social Search Based on the User's Social Network. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management (CIKM '09)*. ACM, New York, NY, USA, 1227–1236. DOI :
<http://dx.doi.org/10.1145/1645953.1646109>
 12. Zhicong Cheng, Bin Gao, and Tie-Yan Liu. 2010. Actively predicting diverse search intent from user browsing behaviors. In *Proceedings of the 19th international conference on World wide web*. ACM, 221–230.
 13. Anind K. Dey, Gerd Kortuem, David R. Morse, and Albrecht Schmidt. 2001. Situated Interaction and Context-Aware Computing. *Personal Ubiquitous Comput.* 5, 1 (Jan. 2001), 1–3. DOI :
<http://dx.doi.org/10.1007/PL00000013>
 14. Joan Morris DiMicco and Walter Bender. 2004. Second messenger: increasing the visibility of minority viewpoints with a face-to-face collaboration tool. In *Proceedings of the 9th international conference on Intelligent user interfaces*. ACM, 232–234.
 15. Joan Morris DiMicco, Anna Pandolfo, and Walter Bender. 2004. Influencing group participation with a shared display. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*. ACM, 614–623.
 16. Desmond Elliott and Joemon M. Jose. 2009. A Proactive Personalised Retrieval System. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management (CIKM '09)*. ACM, New York, NY, USA, 1935–1938. DOI :
<http://dx.doi.org/10.1145/1645953.1646269>
 17. Greg Guest. 2014. *Public health research methods*. Sage Publications.
 18. Ramanathan Guha, Vineet Gupta, Vivek Raghunathan, and Ramakrishnan Srikant. 2015. User Modeling for a Personal Assistant. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15)*. ACM, New York, NY, USA, 275–284. DOI :
<http://dx.doi.org/10.1145/2684822.2685309>
 19. Qi Guo and Eugene Agichtein. 2010. Ready to Buy or Just Browsing?: Detecting Web Searcher Goals from Interaction Data. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '10)*. ACM, New York, NY, USA, 130–137. DOI :
<http://dx.doi.org/10.1145/1835449.1835473>
 20. Monika Henzinger, Bay-Wei Chang, Brian Milch, and Sergey Brin. 2005. Query-Free News Search. *World Wide Web* 8, 2 (2005), 101–126. DOI :
<http://dx.doi.org/10.1007/s11280-004-4870-6>
 21. Jonathan Kilgour, Jean Carletta, and Steve Renals. 2010. The Ambient Spotlight: Queryless Desktop Search from Meeting Speech. In *Proceedings of the 2010 International Workshop on Searching Spontaneous Conversational Speech (SSCS '10)*. ACM, New York, NY, USA, 49–52. DOI :
<http://dx.doi.org/10.1145/1878101.1878112>
 22. David Kirsh. 2014. The importance of chance and interactivity in creativity. *Pragmatics & Cognition* 22, 1 (2014), 5–26.

23. Weize Kong, Rui Li, Jie Luo, Aston Zhang, Yi Chang, and James Allan. 2015. Predicting Search Intent Based on Pre-Search Context. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '15)*. ACM, New York, NY, USA, 503–512. DOI : <http://dx.doi.org/10.1145/2766462.2767757>
24. Henry Lieberman. 1995. Letizia: An agent that assists web browsing. In *Proc. IJCAI (1)*. 924–929.
25. Daniel J. Liebling, Paul N. Bennett, and Ryen W. White. 2012. Anticipatory Search: Using Context to Initiate Search. In *Proc. SIGIR*. ACM, 1035–1036.
26. Moira McGregor and John C. Tang. 2017. More to Meetings: Challenges in Using Speech-Based Technology to Support Meetings. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 2208–2220. DOI : <http://dx.doi.org/10.1145/2998181.2998335>
27. Donald McMillan, Antoine Lorient, and Barry Brown. 2015. Repurposing Conversation: Experiments with the Continuous Speech Stream. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3953–3962. DOI : <http://dx.doi.org/10.1145/2702123.2702532>
28. Michael F McTear. 2002. Spoken dialogue technology: enabling the conversational user interface. *ACM Computing Surveys (CSUR)* 34, 1 (2002), 90–169.
29. Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. *CoRR* abs/1301.3781 (2013). <http://arxiv.org/abs/1301.3781>
30. Takashi Nakamura. 2015. The action of looking at a mobile phone display as nonverbal behavior/communication: A theoretical perspective. *Computers in Human Behavior* 43 (2015), 68 – 75. <https://doi.org/10.1016/j.chb.2014.10.042>
31. Matteo Negri, Marco Turchi, José G. C. de Souza, and Daniele Falavigna. 2014. Quality Estimation for Automatic Speech Recognition. In *COLING 2014, 25th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers, August 23-29, 2014, Dublin, Ireland*. 1813–1823. <http://aclweb.org/anthology/C/C14/C14-1171.pdf>
32. Martin Porcheron, Joel E. Fischer, and Sarah Sharples. 2016. Using Mobile Phones in Pub Talk. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 1649–1661. DOI : <http://dx.doi.org/10.1145/2818048.2820014>
33. Pearl Pu, Li Chen, and Rong Hu. 2012. Evaluating recommender systems from the user's perspective: survey of the state of the art. *User Modeling and User-Adapted Interaction* 22, 4 (01 Oct 2012), 317–355. DOI : <http://dx.doi.org/10.1007/s11257-011-9115-7>
34. Paul Resnick and Hal R. Varian. 1997. Recommender Systems. *Commun. ACM* 40, 3 (March 1997), 56–58. DOI : <http://dx.doi.org/10.1145/245108.245121>
35. Bradley Rhodes and Thad Starner. 1996. Remembrance Agent: A continuously running automated information retrieval system. In *The Proceedings of The First International Conference on The Practical Application Of Intelligent Agents and Multi Agent Technology*. 487–495.
36. B. J. Rhodes and P. Maes. 2000. Just-in-time Information Retrieval Agents. *IBM Syst. J.* 39, 3-4 (July 2000), 685–704. DOI : <http://dx.doi.org/10.1147/sj.393.0685>
37. Tuukka Ruotsalo, Giulio Jacucci, Petri Myllymäki, and Samuel Kaski. 2014. Interactive Intent Modeling: Information Discovery Beyond Search. *Commun. ACM* 58, 1 (Dec. 2014), 86–92. DOI : <http://dx.doi.org/10.1145/2656334>
38. G. Salton, A. Wong, and C. S. Yang. 1975. A Vector Space Model for Automatic Indexing. *Commun. ACM* 18, 11 (Nov. 1975), 613–620. DOI : <http://dx.doi.org/10.1145/361219.361220>
39. Gianluca Schiavo, Alessandro Cappelletti, Eleonora Mencarini, Oliviero Stock, and Massimo Zancanaro. 2014. Overt or subtle? Supporting group conversations with automatically targeted directives. In *Proceedings of the 19th international conference on Intelligent User Interfaces*. ACM, 225–234.
40. Yang Shi, Yang Wang, Ye Qi, John Chen, Xiaoyao Xu, and Kwan-Liu Ma. 2017. IdeaWall: Improving Creative Collaboration Through Combinatorial Visual Stimuli. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 594–603. DOI : <http://dx.doi.org/10.1145/2998181.2998208>
41. Sosuke Shiga, Hideo Joho, Roi Blanco, Johanne R. Trippas, and Mark Sanderson. 2017. Modelling Information Needs in Collaborative Search Conversations. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '17)*. ACM, New York, NY, USA, 715–724. DOI : <http://dx.doi.org/10.1145/3077136.3080787>
42. Yang Song and Qi Guo. 2016. Query-Less: Predicting Task Repetition for NextGen Proactive Search and Recommendation Engines. In *Proceedings of the 25th International Conference on World Wide Web (WWW '16)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 543–553. DOI : <http://dx.doi.org/10.1145/2872427.2883020>

43. Jaime Teevan, Susan T. Dumais, and Eric Horvitz. 2005. Personalizing Search via Automated Analysis of Interests and Activities. In *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '05)*. ACM, New York, NY, USA, 449–456. DOI : <http://dx.doi.org/10.1145/1076034.1076111>
44. Jaime Teevan, Susan T. Dumais, and Eric Horvitz. 2010. Potential for Personalization. *ACM Trans. Comput.-Hum. Interact.* 17, 1, Article 4 (April 2010), 31 pages. DOI : <http://dx.doi.org/10.1145/1721831.1721835>
45. Jaime Teevan, Amy Karlson, Shahriyar Amini, A. J. Bernheim Brush, and John Krumm. 2011. Understanding the Importance of Location, Time, and People in Mobile Local Search Behavior. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, USA, 77–80. DOI : <http://dx.doi.org/10.1145/2037373.2037386>
46. Tung Vuong, Giulio Jacucci, and Tuukka Ruotsalo. 2017a. Proactive Information Retrieval via Screen Surveillance. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '17)*. ACM, New York, NY, USA, 1313–1316. DOI : <http://dx.doi.org/10.1145/3077136.3084151>
47. Tung Vuong, Giulio Jacucci, and Tuukka Ruotsalo. 2017b. Watching Inside the Screen: Digital Activity Monitoring for Task Recognition and Proactive Information Retrieval. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 109 (Sept. 2017), 23 pages. DOI : <http://dx.doi.org/10.1145/3130974>
48. Ryen W. White and Resa A. Roth. 2009. *Exploratory Search: Beyond the Query-Response Paradigm*. Morgan & Claypool Publishers, [San Rafael, Calif.].
49. Biao Xiang, Daxin Jiang, Jian Pei, Xiaohui Sun, Enhong Chen, and Hang Li. 2010. Context-aware ranking in web search. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 451–458.